

An Evaluation of Data and Data Collection Processes Regarding Biased-Based Policing

Utah Commission on Criminal and Juvenile Justice
September 2003

During the 2001 Legislative Session, House Bill 101 was passed which focused on biased policing. The Utah Department of Public Safety (DPS) was required to collect data from local law enforcement agencies and provide it to the Utah Commission on Criminal and Juvenile Justice (CCJJ) for evaluation purposes. The data collected included the race and gender of the officer, the race of the stoppee, and the purpose of the stop.

Specifically, CCJJ was tasked with evaluating the data, evaluating the effectiveness of the data collection process, and reporting and making recommendations to the Legislature. This report fulfills CCJJ's responsibility regarding data collection for biased policing for 2003.

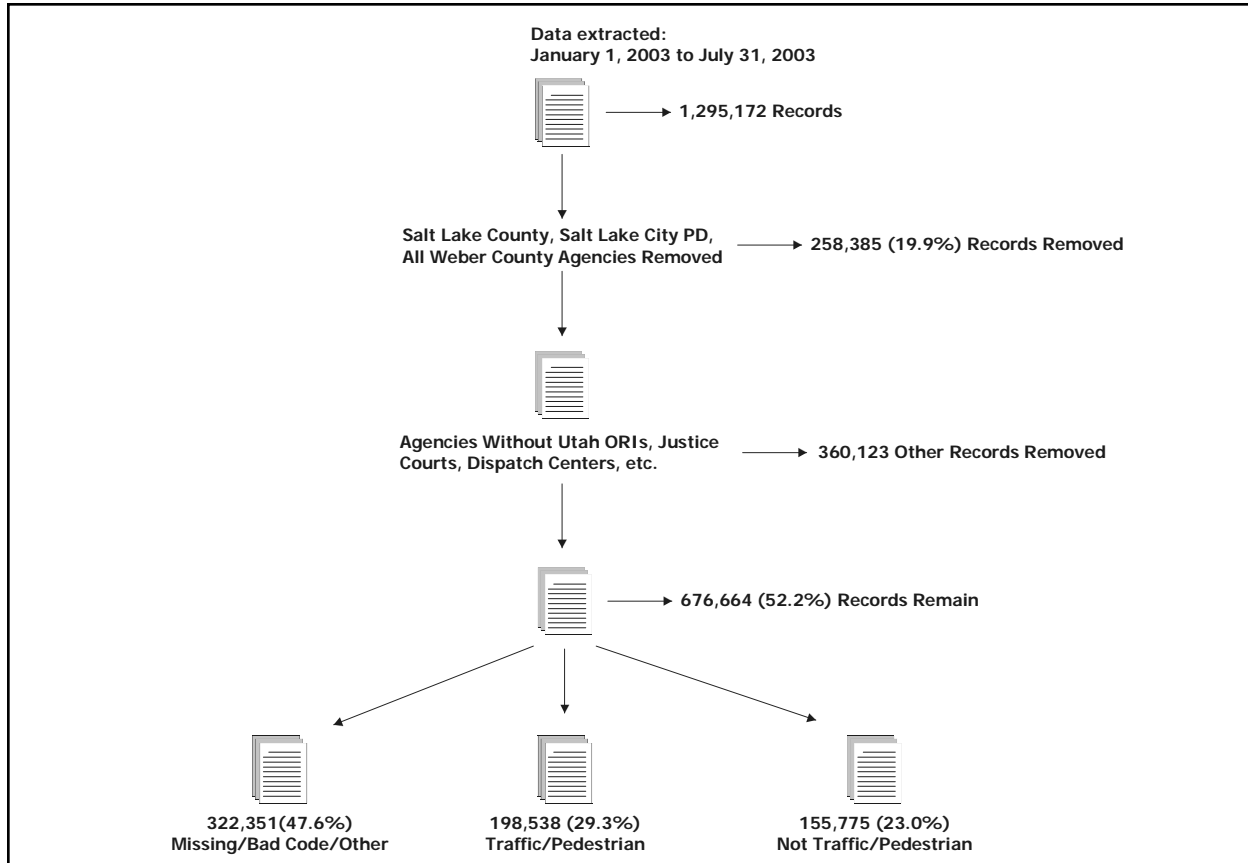
Evaluating the Data

It is important to understand how the process for data collection for biased policing was established by DPS. Each law enforcement agency in Utah was to submit to DPS a list of officers, their identification numbers, race, and gender. In doing so, when an officer in the field enters their identification number for a driver's license transaction, their race and gender is automatically entered on the screen. In theory, the only additional step officers need to take when reviewing a driver's license is entering the reason or purpose for pulling the driver's information.

Very general purpose categories were established, allowing simple entry for law enforcement and a method for CCJJ to differentiate between traffic activities and non-traffic activities, such as investigations. At the beginning of this effort, it was deemed important that researchers be able to look only at those transactions associated with traffic or pedestrian related situations. For example, if a detective were looking for a suspect in an investigation who was described as an Hispanic male, he would likely run several transactions specifically related to Hispanics, and may, therefore, look like he is engaged in biased policing. These situations had to be identified and excluded. Additionally, statute requires the collection of this data only for traffic and pedestrian stops.

As of mid-year 2003, data collected in this effort are scarce. Several of the largest law enforcement agencies are not yet reporting data. This includes the Salt Lake County Sheriff's Office, the Salt Lake City Police Department, and all agencies in Weber County. All of these agencies use the Versa-Term records management system, which required a \$50,000 upgrade in order to report the biased-based policing data. Budget constraints would not allow these agencies to begin reporting data until the end of 2003 or beginning of 2004.

Figure 1: Biased-Based Policing Data Evaluation



CCJJ extracted the records relevant to biased policing between January 1, 2003 and July 31, 2003 (*Figure 1* depicts data evaluation issues). These records contain identifying information for the person conducting the transaction, the employing agency, his/her race, his/her gender, the purpose of the transaction, and the race of the individual being reviewed. During this seven month period, 1,295,172 transactions occurred. However, the previously noted agencies were included within this total. Once these agencies (Salt Lake County Sheriff, Salt Lake City PD, and the Weber County agencies) were removed, 1,036,787 records remained. The dataset also included 360,123 transactions run by

agencies outside of Utah, Justice Courts, dispatch centers, etc. These were also removed, leaving 676,664 records. For the remainder of this analysis, the term "requester" will refer to the individual who ran the transaction, and the term "subject" will refer to the individual who's record was being accessed.

In order to conduct the analysis of traffic/pedestrian stops, it was critical to know why each transaction was run. When reviewing the purpose identified for each transaction, it was found that nearly half of the records (47.6%) were missing the purpose code, had an erroneous purpose code listed, or were listed as "other". This accounted for 322,351 of the records. Those

records with an identifiable purpose code, besides “other,” were evenly split between traffic/pedestrian (198,538 or 29.3%) and investigations/bookings/vehicle inspections (155,775 or 23.0%).

The missing purpose codes have an important impact on our ability to examine racial characteristics of the subjects and the requesters. The number of traffic/pedestrian stops, statewide and by individual agency, becomes the denominator in calculating the proportional distribution of race. In essence, it is half of the equation. With half of the purpose codes missing, the number of traffic stops could be as high as 520,889 (198,538 traffic stops + 322,351 missing/bad codes/other) or as low as 198,538. In reality, the true number likely lies somewhere in between. With such an inflammatory issue as racial profiling, it is dangerous to estimate the percentage of missing data that might be traffic related.

To demonstrate the impact this could have on racial proportionality, if we knew 20,000 of those traffic stops were racial minorities, the percentage of stops of minorities could range from 10.1% (20,000/198,538) to 3.8% (20,000/520,889). The first figure may hint at biased policing, while the second figure may not.

To further complicate matters, *Figure 2* depicts the demographic information for both the requesters and the subjects. Requester demographic information is provided by the law enforcement agency, while the subject demo-

Figure 2: Demographic Information

Requester Race		
	<i>n</i>	%
Non-Minority	274,935	40.6%
Minority	11,958	1.8%
Unknown	82,052	12.1%
Missing Data	307,719	45.5%

Requester Sex		
	<i>n</i>	%
Missing	222,663	32.9%
Female	76,492	11.3%
Male	377,509	55.8%

Subject Race		
	<i>n</i>	%
Non-Minority	222,846	32.9%
Minority	43,811	6.5%
Unknown	401,999	59.4%
Missing Data	8,008	1.2%

graphic information is extracted directly from the Drivers License Database wherein drivers optionally self-identify their race.

Looking first at the race of the requester, in 57.6% (389,771) cases, the race of the requester was not known. In 32.9% (222,663) of the cases, the sex of the requester was not known. This is information that is supposed to be provided by local law enforcement agencies.

However, even more problematic is the amount of subject race data that was not available. As noted previously, subject race is provided by the subject on their Drivers License. Provision of this data is optional. Looking at *Figure 2*, in 60.6% (410,007) of the records, the subject's race was not known. This is very problematic for a variety of reasons. First, subject race would be the cornerstone of any assessment

of biased policing. Without knowing the race of the individual involved in a traffic or pedestrian stop, it is difficult, at best, to establish correlation between race and the stop. Second, it is problematic in that there is little we can do to improve the situation, short of requiring drivers to report their race when applying for a Drivers License.

It is possible the subject race numbers could improve over time. However, it is important to realize new licensees have five years before being required to renew. In short, although one can be optimistic these figures will improve, one should be prepared for a relatively gradual improvement, if at all.

Summary of Data Evaluation

As a whole, the lack of data on various levels makes analysis of biased policing very problematic. As will be discussed in the following sections, identifying racial bias in policing generally requires the examination of multiple data elements, beyond what is required in Utah statute. In Utah, we have been provided only one useful data element, the race of the subject, and that element is unknown in over 60.0% of the cases.

As the data was examined, the problems avalanched. We began with nearly 1.3 million records, which was pared down to just over 1 million records because many large agencies are unable to report data until late 2003. Narrowing the data to law enforcement agencies only, left nearly 700,000 records. Of these, nearly half were unusable because we were unable to sift the traffic-related transactions from the non-traffic transactions (i.e. the purpose code was missing or bad). Finally, the most important piece of data for this analysis, race of the subject, was not known in over 60.0% of the cases.

With the data collected during 2003, CCJJ stands unable to calculate the proportion of minorities stopped in a specific community because 1) in half of the cases, we don't know whether or not the transaction was related to a traffic stop; and 2) in over half of the cases, we don't know the race of the individual involved in the incident.

Evaluating the Effectiveness of the Data Collection Process

Two primary issues are discussed in this section of the report. The first is whether the data collection system, as currently constituted, is effective in collecting the data it was established to collect. The second is whether sufficient data elements are being collected to determine

whether biased-based policing is occurring in Utah.

Current Data Collection System

The data collection system established after legislation was passed in 2001 attempts to collect the required data with the least time requirements for law enforcement, in terms of data entry. Once each agency provided DPS a table containing each officer's race and gender, additional data entry time for officers on the street should be minimal. In theory, when the officer makes a stop and enters his/her identification number on the data entry screen, his/her race and gender would pre-fill the screen, requiring no data entry. The only additional step required by the officer would be selecting the purpose for the record look-up, which allows CCJJ to differentiate between investigational lookups and traffic stops. The race of the driver is to be automatically extracted from the Drivers License Database, based upon the drivers license number.

DPS completed the required programming; however, there were a few outstanding problems with this data collection approach.

One of the first obstacles in data collection was a realization that technical difficulties would not allow Salt Lake City, Salt Lake County, and all Weber County agencies from using the system developed by DPS. These agencies all use the VersaTerm records management system,

which would tie into the system developed by DPS. After discussions with VersaTerm, it was determined a \$50,000 expenditure would be needed to allow these agencies to collect the mandated data immediately. A secondary option was to wait 18 months for a "service release" of the VersaTerm product that would include the mandated data. This option would not have the \$50,000 cost, but would not begin collecting data until late 2003 or early 2004. This second option was adopted by the agencies in question.

A second, but equally important, obstacle is the absence of racial data on Utah drivers. The race of the driver is central to any analysis of biased-based policing. This data is collected when drivers apply for or renew their drivers license. Reporting race on the drivers license is not mandatory. Looking at the data for the first half of 2003, nearly two-thirds of the drivers license transactions were missing the race of the driver. This is primarily because drivers are not or have not yet self-identified their race for their licenses.

Although the number of drivers reporting their race on their licenses may improve over time, there is no guarantee this will be done in sufficient number to make analysis reliable. Other approaches could be enlisted, such as making race identification on drivers licenses mandatory or by having law enforcement officers note the driver's race at the time of the stop. Each of these approaches has significant drawbacks.

Additionally, without knowing the purpose for the license look-up, CCJJ is unable to exclude those transactions that were not traffic/pedestrian related. The statute requires CCJJ to specifically review traffic and pedestrian stops. In reality, drivers licenses are examined for numerous law enforcement purposes outside of traffic/pedestrian incidents. CCJJ worked with DPS and the Chiefs and Sheriffs Associations to develop a short list of purposes that would help CCJJ differentiate traffic/pedestrian transactions from those non-traffic/pedestrian transactions. According to the 2003 data, the purpose for the drivers license look-up was missing for half of the cases, leaving CCJJ unable to determine whether or not they were traffic related.

Finally, the race and sex of the officer is required for each traffic/pedestrian transaction. Although this data is not extremely relevant to the analysis of biased-based policing, it was also missing or unknown in most cases. This information was to be provided by agencies to DPS, then automatically entered as the officer ran drivers license transactions.

The collection of data regarding biased-policing is a significant change in traditional law enforcement responsibility. It is difficult to determine whether the large amount of missing/unknown data is due to technical problems or to implementation problems. Often, large changes in operations across many organizations take time to fully implement. In time, law enforce-

ment officers may become accustomed to entering the required data, and drivers may become more willing to report their race on their drivers license. However, at this time, looking at each of the required data elements (race of requester, sex of requester, purpose for stop, and race of subject), over half of the data is either unknown or missing.

Evaluation of Data Elements

Perhaps the greatest concern regarding Utah's efforts to study biased policing is the lack of sufficient information to draw conclusions. This is an issue of needing additional data elements to more fully understand police behavior.

The data elements currently required by statute provide an incredibly limited perspective, and it is dangerous to use these elements to draw conclusions regarding biased policing. Upon close examination, the race of the driver is the only data element pertinent to this analysis. Collecting and reviewing the race of the officer is based upon an unproven conclusion that only non-minority officers engage in the practice of racial profiling. The gender of the officer is also not very relevant to the analysis.

Racial profiling, or biased policing, is most likely to occur in low visibility/high discretionary circumstances. In order to seriously analyze the issue, data must be collected that exposes information about these types of situations. The Police Executive Research Forum (PERF),

has published extensively on data collection and analysis of data regarding racially biased policing.¹ The following is a listing and explanation of various additional data elements that PERF suggests could be collected to better understand the issue:

Date/Time: It is important to know not only the date of the stop, but also the time. For metropolitan areas like Salt Lake City, daytime populations may be quite different due to the inflow of workers from suburban areas. This can impact the racial composition of the city which is our baseline for comparison.

Address: In larger cities, minorities are not always evenly distributed geographically. Officers patrolling dominantly minority areas of the city may appear to be profiling if their numbers are compared against the entire city's population demographic. Thus, it is critical to know where the stop occurred and where the officer has been assigned to patrol.

Race/Ethnicity: This is the key data element. Researchers must know the race/ethnicity of the driver. Whether it is important that this information is self-reported or identified by the officer is open for debate.

Resident of Jurisdiction: This also helps us understand our baseline measure, which is the demographic of the city in which the stop is made. If we find many stops are of people who

do not reside in the jurisdiction, it would be inaccurate to compare the demographic of those stopped to the demographic of the city.

Stop Reactive or Self-Initiated:

Officer discretion is important to understand. Stops where the officer was called to service are not discretionary, and should not be included in the assessment of racial profiling. Profiling analysis should focus on those situations where the officer has a high amount of discretion, e.g. self-initiated stops.

Vehicle Code Violation: This data element goes deeper into the purpose for the stop. Rather than just identifying the stop as traffic or non-traffic, it is suggested data be collected on whether the traffic violation was, for example, a failure to stop or a lane change violation. Although on the surface this may not seem important, it does help arrive at the issue of discretion. For example, officers have a lot of discretion when making a decision to stop a person for failure to signal while changing lanes, but they have little discretion when making a decision to stop a person for running a red light.

Disposition: It is critical to know the outcome of each stop. Was a citation issued or was the driver simply given a warning? Again, this is an issue of officer discretion. Officers who are profiling, may be more likely to make traffic/pedestrian stops without issuing a citation.

Was the Driver Visible? This is simply an issue of fairness. If the officer stops a person without seeing the driver, it makes little sense in evaluating the stop in the context of racial profiling. However, one could argue an officer who is profiling could simply note that he/she did not see the driver prior to the stop in order to conceal the profiling behavior.

Was there a Search? This could be a search of an individual or a search of a vehicle. It is an issue of officer discretion and opportunity for harassment if there is not a legitimate reason to conduct the search.

What was searched? This helps clarify the previous question. The data element would include vehicle, personal effects, driver, and passenger.

Authority to Search: Under what authority was the search conducted? This would include responses such as consent, reasonable suspicion, probable cause, or an item was in plain view.

Result of the Search: This question records whether or not the search resulted in a positive or negative finding.

The search related questions can help researchers understand whether or not minorities are more likely to have their vehicles searched, under what conditions they are searched, and what are the results of the search.

These are a handful of data elements suggested for collection when assessing biased policing. In Utah, we are collecting only two or three of these data elements. This puts us at a disadvantage in our attempt to comment on the status of racial profiling.

Looking at the additional data elements helps us understand some of the difficulties in drawing conclusions from aggregate data. It is not a viable prospect with only one data element. Take, for example, the concept of benchmarking. Benchmarking creates our baseline for comparison. In Utah, we would need to use the racial construction of the population for the agency in question. For example, in Salt Lake City, we would use the racial composition for the city compared to the race of individuals stopped.

This is problematic for various reasons. First, a large city may have concentrations of minorities living in specific areas. Officers assigned to patrol these areas may appear racially biased if we are comparing the demographics of their stops to the demographic of the entire city. Second, large cities' racial demographic may change throughout the day during the week, as commuters enter and exit the city. Additionally, officers may be called for service at higher rates in certain areas of the city that reflects a specific demographic. Officers have little discretion in their decision whether or not to respond to the call.

By collected the additional data elements, many of these problems could be confronted. Without the additional data, conclusions about an agency could be drawn that are erroneous. Drawing improper conclusions on such a high-profile issue as biased policing could have detrimental impacts both on the agency and the community.

However, it is also important to understand the impact collecting the additional information could have on law enforcement and community safety. Entering many elements for each stop (traffic and pedestrian), when aggregated, would take a lot of the officer's time. This is time the officer would generally use to impact public safety. Collecting the information may also have adverse effects on the morale within the law enforcement agency. One also must consider the fact that data collection has been incredibly poor for the limited information required by statute. Although collecting the additional data would assist in assessing biased policing, policy makers must determine the costs and benefits in proceeding in that direction.

Utah Methodology

With the few data elements currently required under statute, assuming we had sufficient data reported, CCJJ would be limited to simply comparing the racial demographic of each city with the racial demographic of the individuals stopped by the law enforcement agency of that

city. It is important to keep in mind the previously stated dangers of this type of analysis.

If CCJJ were to proceed with this type of analysis, we could only identify anomalies that exist agency by agency. Hypothetically, CCJJ might find "City A's" population is 8.0% minority, and 15.0% of the individuals stopped in "City A" are minority. This would be an apparent discrepancy, but would not necessarily imply racial profiling is rampant within the city.

Several dangers exist with this type of conclusion. First, an agency may be labeled racist by the community and/or media, when, in fact it is not. Second, and equally dangerous, there could be agencies who have officers involved in biased practices that are not more intimately reviewed because an anomaly did not arise between the racial demographic of the city and the race of the individuals stopped in the city.

Recommendations

Data collected as required by statute is inadequate to conduct analysis of biased policing. This is due both to the missing, statutorily required data and the lack of collection of additional critical data elements.

Collection of the statutorily required data elements may improve over time. DPS may assist this effort by requiring all data elements to be entered by police officers in order for them to

run the transactions. In this scenario, officers would not be allowed to “skip” racial profiling data elements when entering data about the stop.

Collecting the race of the driver, the most critical data element of all, is more problematic. At this time, reporting of race on the drivers license is purely voluntary on the part of the driver. Trends indicate this reporting is improving, but it is difficult to divine how much it will improve over the coming years. It is important to realize this will be a slow process because drivers only have to renew their licenses every five years. One option to assist in collecting this critical data would include making mandatory the reporting of race for drivers licenses. This option, however, would also likely take years to collect and could create consternation in the community.

Those who have been heavily involved in biased policing research would suggest Utah is not collecting sufficient data to draw reliable conclusions. It would be irresponsible, even if we had sufficient data collected, to draw strong conclusions regarding racial profiling using the minimal data elements we are currently collecting. Collecting additional data elements, documented earlier in this report, would assist in drawing better conclusions. However, collecting additional data elements would cost the state in terms of law enforcement time and morale. Policy makers must balance the research needs in this area with the impact it may have on law enforcement and community safety.

A final consideration, is to move beyond research and into solutions. Finding hard facts about racial profiling in aggregate data is difficult and illusive. Given the data elements not being collected and the status of those which are being collected, it is questionable that CCJJ could ever definitively pinpoint profiling behavior. On the outside chance profiling was ascribed to a particular agency, approaches, such as training, corrective action, and community outreach, would be conducted by the agency. Many of these steps might be beneficial to law enforcement agencies and communities without racial profiling ever being discovered. Data currently being collected could continue to be collected, and agencies who have complaints against specific officers could use the data for evaluative purposes. In short, Utah may be further ahead by proceeding with solutions, rather than chasing an answer that we may not be able to find.

Endnote:

¹ “Analyzing/Interpreting Police-Citizen Contact Data Collected to Assess Racial Biased Policing.” Police Executive Research Forum. www.policeforum.org.